

# On the computational complexity of finding hard tautologies

Jan Krajíček\*

Faculty of Mathematics and Physics  
Charles University in Prague

## Abstract

It is well-known (cf. K.-Pudlák[18]) that a polynomial time algorithm finding tautologies hard for a propositional proof system  $P$  exists iff  $P$  is not optimal. Such an algorithm takes  $1^{(k)}$  and outputs a tautology  $\tau_k$  of size at least  $k$  such that  $P$  is not p-bounded on the set of all  $\tau_k$ 's.

We consider two more general search problems involving finding a hard formula, **Cert** and **Find**, motivated by two hypothetical situations: that one can prove that  $\text{NP} \neq \text{coNP}$  and that no optimal proof system exists. In **Cert** one is asked to find a witness that a given non-deterministic circuit with  $k$  inputs does not define  $\text{TAUT} \cap \{0, 1\}^k$ . In **Find**, given  $1^{(k)}$  and a tautology  $\alpha$  of size at most  $k^{c_0}$ , one should output a size  $k$  tautology  $\beta$  that has no size  $k^{c_1}$   $P$ -proof from substitution instances of  $\alpha$ .

We shall prove, assuming the existence of an exponentially hard one-way permutation, that **Cert** cannot be solved by a time  $2^{O(k)}$  algorithm. Using a stronger hypothesis about the proof complexity of Nisan-Wigderson generator we show that both problems **Cert** and **Find** are actually undefined for infinitely many  $k$ . The results are based on interpreting the Nisan-Wigderson generator as a proof system.

A **propositional proof system** in the sense of Cook and Reckhow [8] is a polynomial time relation  $P(x, y)$  such that for a binary string  $\tau$ :

$$\tau \in \text{TAUT} \text{ iff } \exists \pi \in \{0, 1\}^* P(\tau, \pi)$$

where TAUT is the set of propositional tautologies (in DeMorgan language for the definiteness). Any string  $\pi$  for which  $P(\tau, \pi)$  holds is called a  $P$ -**proof** of  $\tau$ . A proof system (tacitly propositional from now on) is **p-bounded** iff there exists a constant  $c \geq 1$  such that the above holds even with the requirement that  $|\pi| \leq |\tau|^c$ . Cook and Reckhow [8] noted that a p-bounded proof system exists

---

\*Supported in part by grant IAA100190902. A part of this research has been done while the author was a visiting fellow of the Isaac Newton Institute in Cambridge (programme *Semantics and Syntax*) in Spring 2012. Also partially affiliated with the Institute of Mathematics of the Academy of Sciences.

iff  $\text{NP} = \text{coNP}$ . Hence proving that no  $p$ -bounded proof system exists would imply  $\text{NP} \neq \text{coNP}$  and thus also  $\text{P} \neq \text{NP}$ . This fact elevated the investigation of lengths of proofs into a fundamental topic of mathematical logic approach to computational complexity.

Strong lower bounds were proved for a variety of proof systems and several different methods for this purpose were invented. Examples of proof systems that appear to be outside of the scope of current methods are the so called **Frege systems**: the usual text-book propositional calculi based on a finite number of axioms schemes and inference rules (only quadratic lower bounds are known for them, cf.[10]). This apparent failure could cause an uninformed reader to dismiss the whole area of proof complexity. However, although we may not be near proving that  $\text{NP} \neq \text{coNP}$ , the lower bounds for weaker proof systems proved so far do have consequences interesting in their own right. For example, a single lower bound for a proof system  $P$  implies time lower bounds for a class  $\text{Alg}(P)$  of SAT algorithms associated with  $P$  and all commonly used SAT algorithms belong to such a class for some  $P$  for which we have an exponential lower bound (cf.[17] and references given there). Another type of consequences can be found in bounded arithmetic: a lengths-of-proofs lower bound for  $P$  often implies the unprovability of a true  $\Pi_1^0$  sentence in a first-order theory  $T_P$  associated with  $P$ . These unprovability arguments do not use Gödel's theorem and the  $\Pi_1^0$  sentences involved have typically a clear combinatorial meaning. And last but not least, any super-polynomial lower bound for  $P$  also implies that  $\text{P} \neq \text{NP}$  is consistent with  $T_P$ . We shall not survey these proof complexity topics in detail here and instead refer the reader to expositions in [11, 12, 14, 23] or in [15, Chpt.27].

In this paper we are interested in the question how hard it is - to be measured in terms of computational complexity here - to find tautologies hard (i.e. requiring long proofs) for a given proof system. Proposing plausible candidates for tautologies hard for Frege systems and for stronger proof systems turned out to be a quite delicate issue. The lack of a variety of good candidates is one of principal obstacles for proving lower bounds for strong systems. Of course, in principle one would be happy to accept a suggestion for such a hard tautology from a friendly oracle. However, the experience with known lower bound proofs shows that it is essential to have explicit formulas with a transparent combinatorial meaning.

In particular, all first super-polynomial lower bounds for proof systems for which we have any such bounds were proved for some sequence of tautologies  $\{\tau_k\}_k$  of size  $|\tau_k| \geq k$  and constructible in polynomial time (or even log space) from  $1^{(k)}$ . This type of sequences of hard tautologies has been considered in [18] and [11, Chpt.14] and it exists for a proof system  $P$  iff  $P$  is not **optimal**, i.e. there exists a proof system  $Q$  that has a super-polynomial speed-up over  $P$  (w.r.t. lengths of proofs) on an infinite set of formulas. It is consistent with the present knowledge, and indeed most researchers seem to conjecture that, that no optimal proof system exist and hence that for each  $P$  a  $p$ -time constructible sequence of hard formulas exist. However, deriving the existence of hard formu-

las from the assumption of non-optimality is not very illuminating: it is a basic proof complexity result that  $P$  cannot admit polynomial size proofs of formulas expressing the soundness of a proof system  $Q$  (these formulas are log space constructible) if  $Q$  has a super-polynomial speed-up over  $P$  on an infinite set of formulas (cf.[11, Chpt.14]). Hence, in a sense, deriving the existence of a polynomial time sequence of hard tautologies from the non-optimality assumption amounts just to restating the assumption in a different terminology. We refer the reader to Beyersdorff-Sadowski[4] for further information and up-to-date references.

We shall consider in this paper two more general search problems in which the task includes a requirement to find a hard tautology. The two problems model in their ways two hypothetical situations: a situation when one can prove  $NP \neq coNP$  (i.e. super-polynomial lower bounds for all proof systems) and a situation when one can prove that no optimal proof system exists by having a uniform method how to construct from a given proof system a stronger one. These two tasks, **Cert** and **Find**, will be defined in Section 1 (in Section 7 we add one more search task **Pair** involving disjoint pairs of NP sets).

We will prove (using the hypothesis of the existence of a hard one-way permutation) that **Cert** cannot be solved by exponential time algorithms and (using a stronger hypothesis about the proof complexity of the Nisan-Wigderson generator) that both **Cert** and **Find** actually cannot be solved at all on infinitely many input lengths. Our primary motivation for this research is to understand what kind of consequences do various - both proven and conjectural - statements about the proof complexity of the Nisan-Wigderson generator have.

The paper is organized as follows. After the motivation and the definition of the search tasks **Cert** and **Find** in Section 1 we review some complexity theory in Section 2 and some proof complexity in Section 3. The hardness results are proved in Sections 4 and 6, respectively (after a proof complexity interlude in Section 5). The paper is concluded by Section 7 considering a related search task for disjoint pairs of sets and a few remarks in Section 8.

We do assume only basic complexity theory and proof complexity (e.g. the well-known relation between reflection principles and simulations). But the reader may still benefit from understanding a wider proof complexity context. In particular, [15, Chpt.27] overviews some fundamental problems of proof complexity and [15, Chpts.29 and 30] survey<sup>1</sup> the theory proof complexity generators (and list relevant literature).

## 1 The search tasks Cert and Find

We are going to consider two search tasks asking us to find formulas with certain properties (and in Section 7 we add one more). Both are more complex

---

<sup>1</sup>One can read these chapters independently of the rest of the book.

than the mere task to construct hard tautologies for a given proof system that was discussed in the introduction. To motivate them we shall describe two thought situations in proof complexity; the search tasks are then abstract (and simplified) versions of those.

First assume that you can prove (i.e. ZFC can) that  $NP \neq coNP$  and thus, in particular, super-polynomial lower bounds for all proof systems. For a proof system  $P$  and a constant  $c \geq 1$  denote by  $LB_P(c)$  the statement

$$\forall 1^{(k)} \exists \tau [|\tau| \geq k \wedge \tau \in TAUT \wedge \forall \pi (|\pi| \leq |\tau|^c \neg P(\tau, \pi))]$$

formalizing a polynomial lower bound for  $P$  with degree  $c$ .

It is easy to see that for any decent proof system (see Section 5 for a formal definition of decency), as long as we can prove some specific polynomial lower bound we can also prove its soundness. The decency assumption allows to extend a proof of a falsifiable formula  $\varphi$  to a proof of 0 and further to a proof of any  $\tau$ , all in polynomial time.

But by a simple application of Gödel's theorem ZFC is not able to prove the soundness of all proof systems. This suggests that we should by proving lower bounds conditioned upon the assumption that  $P$  is indeed a Cook-Reckhow proof system. If  $P$  were not complete we do not need to bother with lower bounds for it, so the interesting clause of the Cook-Reckhow definition that is of interest here is the soundness and we are lead to implications:

$$Ref_P \rightarrow LB_P(c)$$

where  $Ref_P$  is a universal sentence (in the language  $L_{PV}$  of Section 3) formalizing that any formula with a  $P$ -proof must be a tautology.

Now we simplify the situation bit more. Let  $D(x, y)$  be a circuit in  $k$  variables  $x = (x_1, \dots, x_k)$  and  $\ell = k^c$  variables  $y = (y_1, \dots, y_\ell)$  which we interpret as the provability relation of a proof system restricted to formulas to size  $k$  and proofs of size at most  $\ell$ . This motivates the following **search task Cert**( $c$ ) defined for any constant  $c \geq 1$ :

- input:  $1^{(k)}$  and a size  $k^{c^2}$  circuit  $D(x, y)$  in  $k$  variables  $x = (x_1, \dots, x_k)$  and  $\ell = k^c$  variables  $y = (y_1, \dots, y_\ell)$
- required output: either a size  $k$  falsifiable formula  $\varphi$  such that  $D(\varphi, y)$  is satisfiable or a size  $k$  tautology  $\tau$  such that  $D(\tau, y)$  is unsatisfiable.

The output of  $Cert(c)$  thus certifies that  $D$  is not a non-deterministic circuit (with input  $x$  and non-deterministic variables  $y$ ) that accepts  $TAUT \cap \{0, 1\}^k$ .

The provability relation of a proof system restricted to size  $k$  formulas and size  $\ell$  proofs can be computed by circuits of size  $(k + \ell)^{O(1)}$ . In the formulation of the problem we have represented the  $O(1)$  constant by  $c$  as well. In addition **Cert** ignores the uniformity of such circuits corresponding to a particular proof system (they can be constructed in log space from  $1^{(k)}, 1^{(\ell)}$ ). This is in line with

the prevailing approach in complexity theory to reduce uniform problems to non-uniform finite combinatorial problems. Finally note that in our simplification we are taking the reflection principle just for the proof lengths corresponding to the lower bound we should witness; this make sense due to the non-uniformity.

The second search task we shall consider is motivated by another thought experiment. Assume that you can prove that no optimal proof system exists and, in fact, that you have a uniform construction that from a proof system  $P$  produces a stronger proof system  $Q(P)$  (i.e. not simulable by  $P$ ). For definiteness, assume that there is one oracle polynomial time machine that for all  $P$  defines  $Q(P)$  when having the oracle for  $P$ . Then we expect to be able to prove

$$Ref_P \rightarrow Ref_{Q(P)}$$

and, most importantly, that it is stronger

$$Ref_P \rightarrow \forall 1^{(k)} \forall \pi (|\pi| \leq k^c) \neg P(\text{ref}_{Q(P)}^k, \pi)$$

where  $\text{ref}_{Q(P)}^k$  is a size  $k^{O(1)}$  tautology formalizing the soundness of  $Q(P)$  w.r.t. all proof of size at most  $k$  (we assume for simplicity that a proof is always at least as long as the formula it proves so one parameter suffices). See a similar formula in (12) in the proof of Lemma 5.3.

Any decent proof system can simulate  $Q(P)$  if it can use  $\text{ref}_{Q(P)}^k$  as extra axioms (see Section 5). In the following problem  $\alpha$  represents a bit more<sup>2</sup> generally any extra axiom.

Let  $P$  be a proof system and  $c_1 \geq c_0 \geq 1$  be constants. Consider the following promise **computational task Find**( $P, c_1, c_0$ ):

- input:  $1^{(k)}$  and a tautology  $\alpha$  such that  $|\alpha| \leq k^{c_0}$
- required output: any size  $k$  tautology  $\beta$  that has no proof in proof system  $P + \alpha$ ,  $P$  augmented by  $\alpha$  as an extra axiom scheme<sup>3</sup>, of size less than  $k^{c_1}$ .

The requirement that the size of  $\beta$  is exactly  $k$  is just for a technical convenience; we could allow any interval  $[k^{\Omega(1)}, k^{O(1)}]$  instead.

## 2 Computational complexity preliminaries

Let  $n \rightarrow m = m(n)$  be an injective function such that  $m(n) > n$  and let  $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$  be a Boolean function. The **Nisan-Wigderson generator**  $NW_{A,f} : \{0, 1\}^n \rightarrow \{0, 1\}^m$  is defined using the notion of a design. A  $(d, \ell)$ -design on  $[n]$  is a set system  $A = \{J_i \subseteq [n]\}_{i \in [m]}$  on  $[n] = \{1, \dots, n\}$  such that:

<sup>2</sup>Really just a bit: for decent  $P$  adding  $\alpha$  is equivalent to adding the reflection principles for  $P + \alpha$ .

<sup>3</sup>The proof system  $P + \alpha$  will be defined in Section 5.

- $|J_i| = \ell$ , for all  $i$ ,
- $|J_i \cap J_j| \leq d$ , for all  $i \neq j$ .

The  $i$ -th bit of  $NW_{A,f}(x)$  is computed by  $f_\ell : \{0,1\}^\ell \rightarrow \{0,1\}$  from the  $\ell$ -bit string  $x(J_i) := x_{j_1} \dots x_{j_\ell}$ , where

$$J_i = \{j_1 < \dots < j_\ell\}$$

and  $f_\ell$  is the restriction of  $f$  to  $\{0,1\}^\ell$ . In the future the parameter  $\ell$  will be determined by  $n$  and we shall denote the restriction  $f$  as well. Nisan and Wigderson[21] showed that there are such designs for a wide range of parameters  $n, m, d, \ell$  and that one can construct them uniformly and feasibly. In particular, we can fix the parameters as follows:

$$\ell := n^{1/3} \quad \text{and} \quad m := 2^{n^\delta} \quad \text{and} \quad d := \log(m), \quad (1)$$

where  $1/3 \geq \delta > 0$  is arbitrary. We shall thus assume that fixing  $n$  and  $\delta$  fixes the other parameters and also some set system  $A_n$  constructed from  $1^{(n)}, 1^{(m)}$  in time  $m^{O(1)}$  and with parameters meeting the requirements. In fact, we need that

$$J_i \text{ is computable from } i \text{ and } 1^{(n)} \text{ in polynomial time.} \quad (2)$$

The design from [21, L.2.5] has this property.

In our construction the function  $f$  will be  $\text{NP} \cap \text{coNP}$ . By this we mean that it is the characteristic function of a language in  $\text{NP} \cap \text{coNP}$ . Hence  $f$  is defined by two NP predicates

$$\exists y(|y| \leq |u|^c \wedge F_0(u, y)) \quad \text{and} \quad \exists y(|y| \leq |u|^c \wedge F_1(u, y)) \quad (3)$$

with  $F_0$  and  $F_1$  polynomial-time relations and  $c$  a constant such that

$$f(u) = a \quad \text{iff} \quad \exists y(|y| \leq |u|^c \wedge F_a(u, y)) \quad (4)$$

for  $a = 0, 1$ . Any string  $y$  witnessing the existential quantifier will be called a witness for  $f(u)$ .

We shall use results from [16] and those do assume that  $f$  has unique witnesses, meaning that for each  $u$  there is exactly one witness for  $f(u)$ . A natural source of  $\text{NP} \cap \text{coNP}$  functions with unique witnesses are hard bits of one-way permutations. That is, for a polynomial time (and intended to be one-way) permutation  $h : \{0,1\}^* \rightarrow \{0,1\}^*$  we have

$$f(u) := B(h^{(-1)}(u)) \quad (5)$$

where  $B(v)$  is a hard bit predicate for  $h$ .

The hardness of one-way permutations is measured as follows. A polynomial time permutation  $h$  is defined to be  $\epsilon(\ell)$  **one way with security parameter**  $t(\ell)$  iff for all  $\ell$  and any circuit  $D$  with  $\ell$  inputs and of size at most  $t(\ell)$  it holds:

$$\text{Prob}_{v \in \{0,1\}^\ell} [D(h(v)) = v] \leq \epsilon(\ell) .$$

Using the Goldreich-Levin theorem we may assume that such a permutation  $h$  has a hard bit function  $B(v)$ . The details can be found in Goldreich [9].

Our construction needs to assume that  $f$  is hard in the sense of Nisan and Wigderson [21]. They define  $f$  to be  $(\epsilon(\ell), S(\ell))$ -hard if for every  $\ell$  and every circuit  $C$  with  $\ell$  inputs and of size at most  $S(\ell)$  it holds:

$$\text{Prob}_{u \in \{0,1\}^\ell} [C(u) = f(u)] < 1/2 + \epsilon/2 .$$

They define then the **hardness of  $f$** , denoted  $H_f(\ell)$ , to be the maximal  $S$  such that the function is  $(1/S, S)$ -hard. This simplification makes sense when  $\epsilon$  has the rate about  $m^{-O(1)}$  as in Nisan and Wigderson [21].

In the proof complexity situations studied in [16] the parameter  $S$  plays the main role, with  $\epsilon$  being primarily of the rate  $\ell^{-O(1)}$ . This corresponds to the fact that in applications of the original Nisan-Wigderson generators  $m$  is usually exponentially large but for various purposes of proof complexity (especially lengths-of-proofs lower bounds) the best choice would be at the opposite end:  $m = n + 1$ . This lead in [16] to keeping  $\epsilon$  and  $S$  separate and using the notion of the approximating hardness (defined there) in place of  $H_f(\ell)$ . In this paper, however, we shall use only those results from [16] where  $m$  is exponentially large as in (1) and thus using the measure  $H_f(\ell)$  suffices here.

A one-way permutation  $h$  with a hard bit  $B$  is **exponentially hard** iff it is  $2^{-\ell^{\Omega(1)}}$  one-way with security parameter  $2^{\ell^{\Omega(1)}}$ . The hardness  $H_f(\ell)$  of  $f$  is then  $2^{\ell^{\Omega(1)}}$  as well. Details of these constructions can be found in Goldreich [9].

We will use in Sections 4 a 7 the hypothesis that an exponentially hard one-way permutation exists instead of the presumably weaker assumption that an  $\text{NP} \cap \text{coNP}$  function  $f$  with unique witnesses and with exponential hardness  $H_f$  exists. The only reason is that the former hypothesis is more familiar than the latter one.

### 3 Proof complexity preliminaries

Although the formulation of the search tasks **Cert** and **Find** may not suggests so explicitly this investigation resulted from a research program in proof complexity about the so called proof complexity generators and we shall use some ideas from this theory.

We shall start with a proof complexity conjecture of Razborov[26, Conjecture 2]. Take an arbitrary string  $b \in \{0,1\}^m$  that is outside of the range  $\text{Rng}(NW_{A_n,f})$  of  $NW_{A_n,f}$ . The statement  $b \notin \text{Rng}(NW_{A_n,f})$  is a coNP property of  $b$  and can be expressed by a propositional formula  $\tau(NW_{A_n,f})_b$  in the sense that

$$\tau(NW_{A_n,f})_b \in \text{TAUT} \quad \text{iff} \quad b \notin \text{Rng}(NW_{A_n,f}) .$$

The construction of the propositional translation of the coNP statement is analogous to the usual proof of the NP-completeness of SAT. The details can be found in any of [6, 11, 23, 15]). Note that the size of the formulas is polynomial in

*m.* **Razborov's conjecture** says that these tautologies are hard for Extended Frege system EF for  $NW_{A_n, f}$  defined as above, with  $m = 2^{n^{\Omega(1)}}$  and based on an  $NP \cap coNP$  function  $f$  that is hard on average for P/poly. Pich [22] proved the conjecture for all proof systems admitting feasible interpolation in place of EF.

In [16] we have considered a generalization of the conjecture. We shall recall only one part of that generalization dealing with exponentially large  $m$ ; in the other parts  $m = n + 1$  and they use the notion of approximating hardness of a function mentioned in the previous section.

**Tentative conjecture 3.1 (Part 3 of Statement (S) of [16])**

*Assume  $f$  is an  $NP \cap coNP$  function with unique witnesses that has an exponential Nisan-Wigderson hardness  $H_f(\ell) = 2^{\ell^{\Omega(1)}}$ .*

*Then there is  $\delta > 0$  such that for  $m(n) = 2^{n^\delta}$  and for any infinite NP set  $R$  that has infinitely many elements whose length equals to  $m(n)$  for  $n \geq 1$  it holds:*

$$Rng(NW_{A_n, f}) \cap R \neq \emptyset.$$

Let us observe that Conjecture 3.1 has a proof complexity corollary including Razborov's conjecture.

**Lemma 3.2**

*Let  $P$  be any proof system. Assume that Conjecture 3.1 holds and that the Nisan-Wigderson hardness  $H_f(\ell)$  of an  $NP \cap coNP$  function  $f$  with unique witnesses is  $2^{\ell^{\Omega(1)}}$ .*

*Then there exists  $\delta > 0$  such that for all  $c \geq 1$ , the size of  $P$ -proofs of formulas  $\tau(NW_{A_n, f})_b$  for all large enough  $n$  and all  $b \notin Rng(NW_{A_n, f})$  of size  $|b| = m(n)$  is bigger than  $|\tau(NW_{A_n, f})_b|^c$ .*

**Proof :**

Note that the set  $R$  of all  $b$  of lengths  $m(n)$  for  $n \geq 1$  for which  $\tau(NW_{A_n, f})_b$  has a  $P$ -proof of size at most  $|\tau(NW_{A_n, f})_b|^c$  is in NP.

**q.e.d.**

Now we recall (a part of) the consistency result from [16] concerning Conjecture 3.1. Its technical heart is a lower bound on complexity of functions solving a certain search task associated with  $NW_{A_n, f}$  and that would, in principle, suffice for our purposes here. Using the consistency result itself, however, seems to decrease the number of technicalities one otherwise needs to discuss.

We first recapitulate a few basic definitions. Cook [6] has defined a theory PV whose language  $L_{PV}$  has a name for every polynomial-time algorithm obtained from a few basic algorithms by the composition and by the limited recursion on notation, following Cobham's [5] characterization of polynomial time. The details of the definition of  $L_{PV}$  can be found in [6, 11] but are not important here. In fact, neither is the theory PV itself as we shall work with the true universal



first-order theory of  $\mathbf{N}$  in the language  $L_{PV}$ . We shall denote this theory  $T_{PV}$ , as in [16]. Note that  $T_{PV}$  contains formulas expressing the soundness of all proof systems.

Let  $f$  be an  $NP \cap coNP$  function defined as in (4). Let us abbreviate by  $G(w, z, x, y)$  the open  $L_{PV}$  formula

$$(z_x = 0 \wedge F_0(w(J_x), y)) \vee (z_x = 1 \wedge F_1(w(J_x), y)) \quad (6)$$

where  $J_x$  is from the set system  $A_n$  (polynomial time definable from  $1^{(n)}$  and  $x$ ) and  $F_0, F_1$  are from (3).

We do not have a symbol in  $L_{PV}$  for the function on  $\{0, 1\}^*$  computed for  $n \geq 1$  on  $\{0, 1\}^n$  by  $NW_{A_n, f}$  as it is not a polynomial time function, and the function has to be defined. One possible formalization of the statement  $NW_{A_n, f}(w) = z$  for  $|w| = n$  and  $|z| = m$  is then

$$\forall x \in [m] \exists y (|y| \leq \ell^c) G(w, z, x, y) \quad (7)$$

with  $c$  from (3). Now we are ready to state the result from [16] we shall need.

**Theorem 3.3 (Krajíček[16, Thm.4.2(part 3)])**

*Assume  $f$  is an  $NP \cap coNP$  function with unique witnesses having the Nisan-Wigderson hardness  $H_f(\ell)$  at least  $2^{\ell^{\Omega(1)}}$ .*

*Then there is  $\delta > 0$  such that for any  $NP$  set  $R$  that has infinitely many elements whose length equals to  $m(n)$  for  $n \geq 1$  and defined by  $L_{PV}$  formula  $\exists v (|v| \leq |z|^d) R_0(z, v)$ , with  $R_0$  open, theory  $T_{PV}$  does not prove the universal closure of the formula*

$$A \rightarrow B$$

where  $A$  is the formula with variables  $v, w, z, n, m, \ell$

$$n = |w| \wedge m = |z| \wedge m = 2^{n^\delta} \wedge \ell = n^{1/3} \wedge |v| \leq m^d \wedge R_0(z, v)$$

and  $B$  is the formula

$$\exists x \in [m] \forall y (|y| \leq \ell^c) \neg G(w, z, x, y) .$$

This statement is in [16] derived from a bit finer model-theoretic result.

## 4 The hardness of task Cert

The argument we shall use to derive the hardness of **Cert** applies to a more general situation which we describe now.

By an  $(NP \cap coNP)/poly$  **algorithm** we shall mean two polynomial time predicates  $F_0(x, y, z)$  and  $F_1(x, y, z)$  and a constant  $c \geq 1$  similarly as in (3) but now with an extra argument  $z$  for the non-uniform advice, and a sequence of advice strings  $\{w_k\}_k$  such that  $|w_k| \leq k^c$  (w.l.o.g. we use constant  $c$  also in the bound to the length of advice strings). We shall assume that

$$\forall x, z (|z| \leq |x|^c) [\exists y (|y| \leq |x|^c) F_0(x, y, z)] \oplus [\exists y (|y| \leq |x|^c) F_1(x, y, z)] \quad (8)$$

is valid where  $\oplus$  is the exclusive disjunction. Thus an  $(\text{NP} \cap \text{coNP})/\text{poly}$  algorithm is an  $\text{NP} \cap \text{coNP}$  set of pairs  $(x, z)$  of appropriate lengths augmented by a sequence of advice strings substituted for  $z$ . In our situation it is more natural to talk about algorithms than sets as we shall be looking for "errors they make". We shall denote such an algorithm  $(\mathcal{F}, \{w_k\}_k)$  where  $\mathcal{F}$  is the triple  $(F_0, F_1, c)$  from (8).

For  $L$  a language let us denote by  $L_k$  the truth table of the characteristic function of  $L$  on  $\{0, 1\}^k$ . If  $L \in \text{NE} \cap \text{coNE}$  then the set  $R^L$  of such strings  $\{L_k \mid k \geq 1\}$  is in  $\text{NP}$  and can be defined by an  $L_{PV}$  formula as

$$z \in R^L \text{ iff } \exists v(|v| \leq |z|^d) R_0^L(z, v) \quad (9)$$

with  $R_0^L$  an open formula. Any  $v$  witnessing the existential quantifier for  $z$  will be called a witness for  $z \in R^L$ . Note that  $TAUT \in \text{NE} \cap \text{coNE}$ .

For a language  $L \in \text{NE} \cap \text{coNE}$  and a triple  $\mathcal{F}$  as in (8) define the **search task**  $\text{Err}(L, \mathcal{F})$  as follows:

- input:  $1^{(k)}$ , string  $L_k$  and a witness  $v$  for  $L_k \in R^L$ , and a string  $w$  such that  $|w| \leq k^c$
- required output: a string  $x \in \{0, 1\}^k$  such that  $\mathcal{F}$  using  $w$  as an advice string makes an error on  $x$ :

$$\forall y(|y| \leq |x|^c) [(x \in L_k \wedge \neg F_1(x, y, w)) \vee (x \notin L_k \wedge \neg F_0(x, y, w))] \quad (10)$$

**Theorem 4.1** *Assume that an exponentially hard one-way permutation exists. Let  $L$  be a language such that  $L \in \text{NE} \cap \text{coNE}$ .*

*Then there exists a triple  $\mathcal{F}$  as in (8) such that no deterministic polynomial time algorithm can solve  $\text{Err}(L, \mathcal{F})$  on all inputs for all sufficiently large lengths  $k$ .*

**Proof :**

Assume that language  $L$  satisfies the hypothesis of the theorem and let  $\mathcal{F}$  be any triple as in (8). Assume that  $\mathcal{A}$  is a deterministic polynomial time algorithm that attempts to solve  $\text{Err}(L, \mathcal{F})$  on all inputs for all  $k \geq k_0$ , for some  $k_0 \geq 1$ .

We are going to define a universal  $L_{PV}$  sentence

$$\Psi_{L, \mathcal{F}, \mathcal{A}, k_0}$$

that is true iff  $\mathcal{A}$  solves  $\text{Err}(L, \mathcal{F})$  for all inputs for all  $k \geq k_0$ .

The sentence  $\Psi_{L, \mathcal{F}, \mathcal{A}, k_0}$  is the universal closure of:

$$C \rightarrow D$$

where  $C$  is the formula

$$|z| = 2^k \wedge |v| \leq |z|^d \wedge R_0^L(z, v) \wedge k \geq k_0 \wedge |w| \leq k^c \wedge x = \mathcal{A}(1^{(k)}, z, v, w)$$

with  $\mathcal{A}$  represented by an  $L_{PV}$  function symbol, and  $D$  is the formula

$$|x| = k \wedge \forall y(|y| \leq |x|^c)[(x \in z \wedge \neg F_1(x, y, w)) \vee (x \notin z \wedge \neg F_0(x, y, w))] .$$

The following should be obvious:

**Claim 1:** *Algorithm  $\mathcal{A}$  solves  $\text{Err}(L, \mathcal{F})$  for all inputs for all  $k \geq k_0$  iff the sentence  $\Psi_{L, \mathcal{F}, \mathcal{A}, k_0}$  is true.*

We are going now to define a specific  $(\text{NP} \cap \text{coNP})/\text{poly}$  triple  $\mathcal{F}$  as in (8) such that  $\Psi_{L, \mathcal{F}, \mathcal{A}, k_0}$  will be false for all  $L \in \text{NE} \cap \text{coNE}$ , all polynomial time algorithms  $\mathcal{A}$  and all  $k_0 \geq 1$ .

Let  $h$  be an exponentially hard one-way permutation with a hard bit  $B$ . Hence the function  $f$  from (4) has the exponential hardness  $H_f(\ell) = 2^{-\ell^{\Omega(1)}}$ . The existence of such  $h$  and  $B$  is guaranteed by the hypothesis of the theorem.

Take  $\delta > 0$  provided by Theorem 3.3 and put  $k := n^\delta$ . Using  $NW_{A_n, f} : \{0, 1\}^n \rightarrow \{0, 1\}^m = \{0, 1\}^{2^k}$  define an  $(\text{NP} \cap \text{coNP})/\text{poly}$  triple  $\mathcal{F} = (F_0, F_1, c)$  as in (8) as follows: put  $c := \delta^{-1}$  and for  $x \in \{0, 1\}^k$ ,  $y \in \{0, 1\}^n = \{0, 1\}^{k^c}$ ,  $w$  of size  $|z| = n$  and  $a = 0, 1$  define:

$$F_a(x, y, w) := [h(y) = w(J_x) \wedge B(y) = a] . \quad (11)$$

In other words, on input  $x$  the algorithm computes the  $x$ -th bit of  $NW_{A_n, f}(w)$ .

**Claim 2:** *For no  $L \in \text{NE} \cap \text{coNE}$ , no polynomial time algorithm  $\mathcal{A}$  and no  $k_0 \geq 1$  is the sentence  $\Psi_{L, \mathcal{F}, \mathcal{A}, k_0}$  true.*

To see this note that by substituting term  $\mathcal{A}(1^{(k)}, z, v, w)$  for  $x$  in  $C \rightarrow D$  and quantifying it existentially  $\exists x(x \in [m])$  allows to deduce from  $\Psi_{L, \mathcal{F}, \mathcal{A}, k_0}$  the universal closure of  $A \rightarrow B$  from Theorem 3.3. Hence, by that theorem,  $\Psi_{L, \mathcal{F}, \mathcal{A}, k_0}$  cannot be true.

Claims 1 and 2 imply the theorem.

**q.e.d.**

We remark that the argument can be actually extended to rule out a larger class of algorithms  $\mathcal{A}$ : the so called Student - Teacher interactive computations of [20, 19] (see also [11]).

Let  $(\mathcal{F}, \{w_k\}_k)$  be as above. Define circuits  $D_k(x, y)$  to be (some canonical) circuits with  $k$  inputs  $x$  and  $k^c$  inputs  $y$  that outputs 1 iff  $F_1(x, y, w_k)$  holds. We can choose  $c \geq 1$  large enough so that  $D_k$  has size at most  $k^{c^2}$ .

Given  $1^{(k)}$ , a time  $2^{O(k)}$  algorithm can compute the string  $\text{TAUT}_k$  (as well as the witness for  $\text{TAUT}_k \in R^{\text{TAUT}}$  required in the general formulation of the theorem). Such an algorithm is then polynomial in the size of  $\text{TAUT}_k$ . If  $\tau$  is a solution to **Cert**( $c$ ) on input  $D_k$  then either  $\tau \in \text{TAUT}_k$  and  $\forall y(|y| \leq k^c) \neg F_1(\tau, y, w_k)$  or  $\tau \notin \text{TAUT}_k$  and  $D_k(\tau, y)$  is satisfiable, in which case  $T_{PV}$

implies that  $\forall y(|y| \leq k^c) \neg F_0(\tau, y, w_k)$ . In other words, an algorithm solving **Cert**( $c$ ) solves **Err**( $TAUT, \mathcal{F}$ ) too. This yields the following statement as a corollary to Theorem 4.1.

**Corollary 4.2** *Assume that an exponentially hard one-way permutation exists. There there is  $c \geq 1$  such that no deterministic time  $2^{O(k)}$  algorithm can solve **Cert**( $c$ ) on all lengths  $k \geq 1$ .*

## 5 Proof systems with advice and with extra axioms

The task **Find** was formulated using the provability in a proof system and in this section we develop a technical tool allowing us to move from  $(NP \cap coNP)/poly$  algorithms to proof systems. We shall recall first the notion of a **proof system with advice** as introduced by Cook-K.[7, Def.6.1]. It is defined as the ordinary Cook-Reckhow proof system (cf.the introduction) except that the binary relation  $P(x, y)$  is decidable in polynomial time using an advice string that depends only on the length of  $x$  (the formula). We say that the advice is polynomial iff its length is  $|x|^{O(1)}$ . This concept has some interesting properties; for example, in the classes of these proof systems - with varying bounds on the size of advice strings - there exists an optimal one. We refer the reader to [7, Sec.6] and to subsequent [1, 2, 3] for further information.

Our aim in this section is to link proof systems with polynomial advice with proof systems with extra axioms, as defined below. A sequence of formulas  $\{\alpha_k\}_k$  will be called **p-bounded** iff  $|\alpha_k| \leq k^{O(1)}$  for all  $k$ .

**Definition 5.1** *Let  $P(x, y)$  be an ordinary Cook-Reckhow proof system.*

1. *For a tautology  $\alpha$  the proof system  $P + \alpha$  is defined as follows:  
a string  $\pi$  is a  $(P + \alpha)$ -proof of formula  $\tau$  iff  $\pi$  is a  $P$ -proof of a disjunction of the form*

$$\bigvee_i \neg \alpha'_i \vee \tau$$

*where  $\alpha'_i$  are arbitrary substitution instances of  $\alpha$  obtained by substituting constants and variables for variables.*

2. *For a p-bounded sequence  $\{\alpha_k\}_k$  of tautologies define a string  $\pi$  to be an  $(P + \{\alpha_k\}_k)$ -proof of formula  $\tau$  iff it is a  $(P + \alpha_k)$ -proof of  $\tau$  for  $k = |\tau|$ .*

We allow only substitutions of constants and variables in instances  $\alpha'_i$  in part 1 as that makes sense for all proof systems (e.g. we do not have to discuss various limitations on depth for constant depth Frege systems) and it suffices here. Systems  $(P + \{\alpha_k\}_k)$  are not meant to genuinely formalize the informal notion of proof systems with extra axioms; such systems should not pose restrictions on which extra axioms can be used in proofs of which formulas. We use them

here only as a technical vehicle allowing us to move from proof systems with advice to ordinary proof systems.

Note that while  $P+\alpha$  is a Cook-Reckhow proof system,  $P+\{\alpha_k\}_k$  is generally not. The following lemma is obvious as we can use the sequence  $\{\alpha_k\}_k$  as advice strings to recognize  $(P + \{\alpha_k\}_k)$  - proofs.

**Lemma 5.2** *Let  $P$  be a Cook-Reckhow proof system. For every  $p$ -bounded sequence  $\{\alpha_k\}_k$  of tautologies  $P + \{\alpha_k\}_k$  is a proof system with polynomial advice in the sense of Cook-K.[7].*

In Section 1 we used informally the notion of a decent proof system, meaning a proof system that can perform efficiently a few simple manipulations with proofs. We shall use the formalization of this notion from [17, Sec.2].

In the following  $\text{sat}_k(u, x, v)$  are formulas for  $k \geq 1$  and suitable  $r = k^{O(1)}$  with  $u = (u_1, \dots, u_k)$ ,  $x = (x_1, \dots, x_k)$  and  $v = (v_1, \dots, v_r)$  such that for all  $a, \varphi \in \{0, 1\}^k$  it holds that:

- $\text{sat}_k(a, \varphi, v) \in \text{TAUT}$  iff  $a$  is a truth assignment satisfying formula  $\varphi$ .

The extra variables  $v$  are used to compute the truth value, as in the NP-completeness of SAT.

A proof system (ordinary or with advice)  $P$  is **decent** iff the following tasks can be performed by polynomial time algorithms<sup>4</sup>:

- D1 From a  $P$ -proof  $\pi$  of formula  $\psi(x)$  and a truth assignment  $a$  to variables  $x$  construct a  $P$ -proof of  $\psi(a)$ .
- D2 Given a true sentence  $\psi$  (i.e. no variables) construct its  $P$ -proof.
- D3 Given  $P$ -proofs  $\pi_1$  of  $\psi$  and  $\pi_2$  of  $\psi \rightarrow \eta$  construct a proof of  $\eta$ .
- D4 Given a formula  $\varphi(u_1, \dots, u_n)$  and a  $P$ -proof of formula  $\text{sat}_k(u, \varphi, v)$  with variables  $u, v$  construct a  $P$ -proof of  $\varphi$ .

Conditions D1-3 are easy to verify for many of the usual proof systems (e.g. Frege systems mentioned in the introduction or resolution). The algorithm for condition D4 is defined by induction on the number of connectives in  $\varphi$ , cf.[11, Chpt.9].

**Lemma 5.3** *Let  $Q$  be a proof system with polynomial advice and  $P$  a decent Cook-Reckhow proof system. Then for every constant  $c \geq 1$  there exists a  $p$ -bounded sequence  $\{\alpha_k\}_k$  of tautologies and  $d \geq 1$  such that:*

*Any tautology  $\tau$  having a  $Q$ -proof of size  $\leq |\tau|^c$  has an  $(P + \{\alpha_k\}_k)$ -proof of size  $\leq |\tau|^d$ .*

---

<sup>4</sup>Polynomially bounded functions would suffice for us here but such a weakening would not put more of the usual proof systems into the class and so we just stick with the definition from [17].

**Proof :**

Assume that  $Q$  uses polynomial size advice string  $w_k$  for formulas of size  $k$ . For  $k \geq 1$  denote by  $\text{prov}_Q^{k^c}(x, y, t, s)$  a propositional formula such that:

- $\text{prov}_Q^{k^c}$  has  $k$  atoms  $x = (x_1, \dots, x_k)$  for bits of a formula,  $k^c$  atoms  $y = (y_1, \dots, y_{k^c})$  for bits of a  $Q$ -proof,  $k^c$  atoms  $t = (t_1, \dots, t_{k^c})$  for bits of an advice and  $k^{O(1)}$  atoms  $s = (s_1, \dots, s_{k^{O(1)}})$  for bits of the computation of the truth value of  $Q(x, y)$  with advice  $t$ ,
- For size  $k$  formula  $\varphi$  and a  $k^c$  size strings  $\pi$  and  $w$ :  $\text{prov}_Q^{k^c}(\varphi, \pi, w, s) \in \text{SAT}$  iff  $Q(\varphi, \pi)$  with advice  $w$  is true.

Then take for  $\alpha_k$  the formula with variables  $x, y, z, s, v$

$$\text{prov}_Q^{k^c}(x, y, t/w_k, s) \rightarrow \text{sat}_k(x, z, v) \quad (12)$$

where  $w_k$  is the string used by  $Q$  for size  $k$  formulas. The formula expresses the soundness of  $Q$  and hence it is a tautology. Its total size is  $k^{O(c)}$ .

Let  $\varphi$  be a size  $k$  formula with variables among  $z = (z_1, \dots, z_k)$  and having a size  $\leq k^c$   $Q$ -proof  $\pi$ . Let  $e$  be bits of the computation of  $Q(\varphi, \pi)$  with advice  $w_k$ .

Take the following substitution instance of  $\alpha_k$ :

$$\text{prov}_Q^{k^c}(\varphi, \pi, w_k, e) \rightarrow \text{sat}_k(\varphi, z, v) \quad (13)$$

**Claim:** For some  $d \geq 1$  depending only on  $c \geq 1$  and  $P$  the formula  $\varphi$  has a  $(P + \alpha_k)$ -proof of size  $\leq k^d$ .

We shall use the decency of  $P$ . By the choice of  $\pi$  and  $e$  the sentence  $\text{prov}_Q^{k^c}(\varphi, \pi, w_k, e)$  is true and hence has, by D2, a size  $k^{O(c)}$   $P$ -proof. Then, using D3, use modus ponens to derive in size  $k^{O(c)}$  the formula

$$\text{sat}_k(\varphi, z, v) .$$

D4 then allows to derive in  $P$  the formula  $\varphi$ , in size  $k^{O(1)}$ . The total size of the  $P$ -proof is  $k^{O(c)}$ .

**q.e.d.**

Note that we have not used the decency condition D1 explicitly; it's role is replaced here by the definition of the system  $P + \alpha_k$  which takes as axioms all substitution instances of  $\alpha_k$ .

Formulas  $\alpha_k$  depend not only on  $k$  and  $w_k$  but also on the bound  $k^c$  to the length of  $y$ . This is the reason why we cannot simply say that  $EF + \{\alpha_k\}_k$  simulates  $P$ .

## 6 The undefinability of tasks Cert and Find

We will need the following notion. For a complexity class  $\mathcal{X}$  and a language  $L$  define the property that  $L$  is **infinitely often in**  $\mathcal{X}$ , denoted  $L \in_{i.o.} \mathcal{X}$ , iff there exists  $L' \in \mathcal{X}$  such that

$$L \cap \{0, 1\}^k = L' \cap \{0, 1\}^k$$

for infinitely many lengths  $k$ . Recall the definition of the class  $(NP \cap coNP)/poly$  at the beginning of Section 4.

The following consequence of Conjecture 3.1 was noted at the end of [15, Sec.30.2] and it uses an idea linking the output/input ratio of proof complexity generators with the unprovability of circuit lower bounds due to Razborov [24], quite similar to the reasoning in Razborov-Rudich [27].

### Lemma 6.1

*Assume that Conjecture 3.1 holds and that an exponentially hard one-way permutation exists. Then for every  $L \in NE \cap coNE$ :*

$$L \in_{i.o.} (NP \cap coNP)/poly .$$

*In particular,  $TAUT \in_{i.o.} NP/poly$ .*

### Proof :

Take  $\delta > 0$  from Conjecture 3.1. Put  $k := n^\delta$  and think of a string  $b \in \{0, 1\}^m$  as of the truth-table of the characteristic function of the language  $L$  on inputs of length  $k$ ; denote it  $L_k$  as in Section 4.

For any language  $L$  in  $NE \cap coNE$  the set of strings  $\{L_k \mid k \geq 1\}$  is in NP: the NP witness can collect all  $2^k$  NE witnesses for each variable setting - this will have size  $2^{O(k)}$  - and check their validity.

In particular, if some  $L \in NE \cap coNE$  would determine the truth-tables  $L_k$  for  $k = n^\delta$  such that all but finitely many lie outside the range of  $NW_{A_n, f}$  we would get a contradiction with Conjecture 3.1. Hence we get:

**Claim:** *For infinitely many  $n$ , for  $k = n^\delta$  and  $m = 2^k$ :*

$$L_k \in \{0, 1\}^m \cap Rng(NW_{A_n, f}) .$$

For  $L_k \in Rng(NW_{A_n, f})$  let  $a \in \{0, 1\}^n$  be such that  $NW_{A_n, f}(a) = L_k$ . Then computing  $L$  on  $i \in \{0, 1\}^k$  amounts to computing  $f$  on  $a(J_i)$ . But by the requirement (2) posed on  $A_n$  the set  $a(J_i)$  can be computed from  $i$  and  $a$  (taken as the advice string) in time polynomial in  $n$  and  $f$  is an  $NP \cap coNP$  function.

**q.e.d.**

This lemma has an immediate consequence for problem **Cert**.

**Corollary 6.2** *Assume that Conjecture 3.1 holds and that an exponentially hard one-way permutation exists.*

*Then for some  $c \geq 1$  the task **Cert**( $c$ ) has no solution for infinitely many lengths  $k \geq 1$ .*

We shall derive the same consequence for the task **Find** using the results of Section 5. For a triple  $\mathcal{F}$  as in (8) define a proof system with polynomial advice  $Q(x, y)$  by:

- either  $0 < |w| \leq |x|^c \wedge F_1(x, y, w)$ ,
- or  $w$  is the empty word and  $y$  is a Frege proof of  $x$ ,

thinking of  $w$  as the advice. Now let  $\{w_k\}_k$  such that  $|w_k| \leq k^c$  be a sequence of advice words defining a proof system with advice  $Q$  (there exists at least one such: the sequence of empty strings).

Let  $P$  be any decent Cook-Reckhow proof system and let formulas  $\alpha_k$  and constant  $d \geq 1$  be those provided by Lemma 5.3 for  $c$  from  $\mathcal{F}$ . Assume  $c_0 \geq 1$  is such that  $|\alpha_k| \leq k^{c_0}$  and assume also w.l.o.g. that  $d \geq c_0$ .

Consider the task  $Find(P, c_0, d)$ . A solution for input  $1^{(k)}$  and  $\alpha_k$  is a size  $k$  tautology  $\beta$  and by the choice of  $c_0, d$  it must be that

$$\forall y(|y| \leq k^c) \neg F_1(\beta, y, w_k) .$$

But Conjecture 3.1 implies analogously as above that for the triple  $\mathcal{F}$  coming from  $NW_{A_n, f}$  there will be infinitely many lengths  $k \geq 1$  and strings  $w_k$  for which  $\alpha_k$  is tautology but no such  $\beta$  exists. Hence we have the following statement.

**Theorem 6.3** *Assume that Conjecture 3.1 holds and that an exponentially hard one-way permutation exists.*

*Then for all decent Cook-Reckhow proof systems  $P$  there are constants  $c_1 \geq c_0 \geq 1$  such that the task **Find**( $P, c_0, c_1$ ) has no solution for infinitely many lengths  $k \geq 1$ .*

## 7 Disjoint NP pairs

Let  $(U, V)$  and  $(A, B)$  be two pairs of disjoint subsets of  $\{0, 1\}^*$ . A **reduction** of  $(A, B)$  to  $(U, V)$  is a function  $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$  such that for all  $u$ :

$$u \in A \rightarrow f(u) \in U \quad \wedge \quad u \in B \rightarrow f(u) \in V .$$

It is (non-uniform) **p-reduction** if  $f$  is (non-uniform) p-time.

Disjoint NP pairs appear quite naturally in several places of proof complexity. Most notably, the reflection principle for a proof system just asserts that two NP sets (that of formulas with bounded size proofs and of falsifiable formulas) are disjoint. A particularly elegant form of this observation was found by Razborov [25] in the notion of the canonical pair of a proof system. Shadowing



the relation of the provability of reflection principles to simulations, a similar one exists between the provability of disjointness of such pairs and simulations. We refer the reader to [23] for more background.

Given two pairs of disjoint sets  $(U, V)$  and  $(A, B)$  and a constant  $c \geq 1$  consider the **search task**  $\mathbf{Pair}_{U,V}^{A,B}(c)$ :

- input:  $1^{(k)}$  and a circuit  $C$  with  $k$  inputs, several outputs and of size at most  $k^c$
- required output: a string  $u \in \{0, 1\}^k$  such that

$$u \in A \wedge f(u) \notin U \quad \text{or} \quad u \in B \wedge f(u) \notin V .$$

In other words, the output string  $u$  certifies that circuit  $C$  is not a reduction of  $(A, B)$  to  $(U, V)$  on  $\{0, 1\}^k$ .

Take a triple  $\mathcal{F}$  as in (8) and define  $U$  and  $V$  to be the sets of pairs  $(x, z)$  such that  $|z| \leq |x|^c$  and  $\exists y(|y| \leq |x|^c) F_0(x, y, z)$  or  $\exists y(|y| \leq |x|^c) F_1(x, y, z)$ , respectively.

For a disjoint pair  $A, B$  of sets such that  $A \in \text{NE} \cap \text{coNE}$  take for language  $L$  on  $\{0, 1\}^k$  simply  $A$ . For  $w$  of size  $\leq k^c$  consider circuit  $C_w$  that takes size  $k$  input  $x$  and outputs the pair  $(x, w)$ ; note that  $|C_w| \leq k^{c+1}$  for  $k \gg 0$ . Then a solution to  $\mathbf{Pair}_{U,V}^{A,B}(c+1)$  for input  $1^{(k)}$  and  $C_w$  is also a solution to  $\text{Err}(L, \mathcal{F})$ . Hence Theorem 4.1 implies the following statement.

**Theorem 7.1** *Assume that an exponentially hard one-way permutation exists.*

*Then there are two disjoint NP sets  $U, V$  and  $c \geq 1$  such that for any pair  $A, B$  of disjoint sets such that  $A \in \text{NE} \cap \text{coNE}$  the task  $\mathbf{Pair}_{U,V}^{A,B}(c)$  is not solvable by a deterministic time  $2^{O(k)}$  algorithm.*

The reader familiar with the canonical pairs of proof systems mentioned earlier may note that these sets are in  $E \subseteq \text{NE} \cap \text{coNE}$  and thus the theorem applies to them.

## 8 Concluding remarks

The role of Conjecture 3.1 is rather ambivalent: it implies that  $\text{NP} \neq \text{coNP}$  (Lemma 3.2) but also that  $\text{TAUT} \in_{i.o.} \text{NP}/\text{poly}$  (Lemma 6.1). This is caused by the dual role of the Nisan-Wigderson generator; it is a source of hard tautologies but also a strong proof system. The reader should consider, before dismissing Conjecture 3.1 because of Lemma 6.1, how little contemporary complexity theory understands about the power of non-uniform advice.

It would be interesting to have a variety of candidate combinatorial constructions  $Q(P)$  of a proof system stronger than  $P$ , as discussed in Section 1. At present only the construction of  $iP$ , the implicit  $P$ , from [13] applies to all proof systems and it is consistent with the present knowledge that it indeed

yields stronger proof systems. An indirect plausible construction of  $Q(P)$  may use the relation between proof systems and first-order theories: take theory  $T_P$  corresponding to  $P$ , extend  $T_P$  by  $Con(T_P)$  (or in some other Gödelian fashion) getting  $S$ , and then take for  $Q(P)$  the proof system simulating  $S$  (cf.[18, 11] about  $T_P$  etc.). But it is hardly combinatorially transparent.

## References

- [1] O. Beyersdorff, J. Kobler and S. Muller, Nondeterministic Instance Complexity and Proof Systems with Advice, in: Proc. 3rd International Conference on Language and Automata Theory and Applications (LATA), Springer-Verlag, LNCS 5457, (2009), pp.164-175.
- [2] O. Beyersdorff, J. Kobler and S. Muller, Proof Systems that Take Advice, *Information and Computation*, Vol. **209** (3), (2011), pp.320-332.
- [3] O. Beyersdorff and S. Muller, Does Advice Help to Prove Propositional Tautologies?, in: Proc. 12th International Conference on Theory and Applications of Satisfiability Testing (SAT), Springer-Verlag, LNCS 5584, (2009), pp.65-72.
- [4] O. Beyersdorff and Z. Sadowski, Characterizing the Existence of Optimal Proof Systems and Complete Sets for Promise Classes. in: A.E.Frid, A.Morozov, A.Rybalchenko, K.W.Wagner (Eds.), *Computer Science - Theory and Applications*, Fourth International Computer Science Symposium in Russia, CSR 2009, Novosibirsk, Russia, (August 18-23, 2009). Lecture Notes in Computer Science **5675**, Springer. (2009), pp.47-58.
- [5] A. Cobham, The intrinsic computational difficulty of functions, in : *Proc. Logic, Methodology and Philosophy of Science*, ed. Y. Bar-Hillel, North-Holland, (1965), pp. 24-30.
- [6] S. A.Cook, Feasibly constructive proofs and the propositional calculus, in: *Proc. 7<sup>th</sup> Annual ACM Symp. on Theory of Computing*, (1975), pp. 83-97. ACM Press.
- [7] S. A. Cook, and J. Krajíček, Consequences of the Provability of  $NP \subseteq P/poly$ , *J. of Symbolic Logic*, **72**(4), (2007), pp. 1353-1371.
- [8] S. A. Cook, and Reckhow, The relative efficiency of propositional proof systems, *J. Symbolic Logic*,**44**(1), (1979), pp.36-50.
- [9] O. Goldreich, *Foundations of cryptography*, Vol.1, Cambridge University Press, (2001).
- [10] J. Krajíček, Speed-up for Propositional Frege Systems via Generalizations of Proofs, *Commentationes Mathematicae Universitatis Carolinae*, **30**(1), (1989), pp.137-140.

- [11] J. Krajíček, *Bounded arithmetic, propositional logic, and complexity theory*, Encyclopedia of Mathematics and Its Applications, Vol. **60**, Cambridge University Press, (1995).
- [12] J. Krajíček, On methods for proving lower bounds in propositional logic, in: *Logic and Scientific Methods* Eds. M. L. Dalla Chiara et al., (Vol. 1 of Proc. of the Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence (August 19-25, 1995)), Synthese Library, Vol.259, Kluwer Academic Publ., Dordrecht, (1997), pp.69-83.
- [13] J. Krajíček, Implicit proofs, *J. of Symbolic Logic*, **69(2)**, (2004), pp.387-397.
- [14] J. Krajíček, Proof complexity, in: Laptev, A. (ed.), European congress of mathematics (ECM), Stockholm, Sweden, June 27–July 2, 2004. Zurich: European Mathematical Society, (2005), pp.221-231.
- [15] J. Krajíček, *Forcing with random variables and proof complexity*, London Mathematical Society Lecture Notes Series, Vol. **382**, Cambridge University Press, (2011).
- [16] J. Krajíček, On the proof complexity of the Nisan-Wigderson generator based on a hard  $\text{NP} \cap \text{coNP}$  function, *J. Mathematical Logic*, Vol.**11(1)**, (2011), pp.11-27.
- [17] J. Krajíček, A note on SAT algorithms and proof complexity, *Information Processing Letters*, **112**, (2012), pp.490-493.
- [18] J. Krajíček and P. Pudlák, Propositional proof systems, the consistency of first order theories and the complexity of computations, *J. Symbolic Logic*, **54(3)**, (1989), pp.1063-1079.
- [19] J. Krajíček, P. Pudlák and J. Sgall, Interactive Computations of Optimal Solutions, in: B. Rovan (ed.): *Mathematical Foundations of Computer Science* (B. Bystrica, August '90), Lecture Notes in Computer Science 452, Springer-Verlag, (1990), pp. 48-60.
- [20] J. Krajíček, P. Pudlák and G. Takeuti, Bounded arithmetic and the polynomial hierarchy, *Annals of Pure and Applied Logic*, **52**, (1991), pp.143–153.
- [21] N. Nisan and A. Wigderson, Hardness vs. randomness, *J. Comput. System Sci.*, Vol.**49**, (1994), pp.149–167.
- [22] J. Pich, Nisan-Wigderson generators in proof systems with forms of interpolation, *Mathematical Logic Quarterly*, **57(3)**, (2011), pp.379-383.
- [23] P. Pudlák, The lengths of proofs, in: Handbook of Proof Theory, S.R. Buss ed., Elsevier, (1998), pp.547-637.

- [24] A. A. Razborov, Unprovability of lower bounds on the circuit size in certain fragments of bounded arithmetic, *Izvestiya of the R.A.N.*, **59(1)**, (1995), pp.201-224.
- [25] A. A. Razborov, On Provably Disjoint NP-pairs, unpublished manuscript, (1994).
- [26] A. A. Razborov, Pseudorandom generators hard for  $k$ -DNF resolution and polynomial calculus resolution, preprint, (May'03).
- [27] A. A. Razborov and S. Rudich, Natural proofs, *J. of Computer and Systems Science*, **55(1)**, (1997), pp.24-35.

**Mailing address:**

Department of Algebra  
 Faculty of Mathematics and Physics  
 Charles University  
 Sokolovská 83, Prague 8, CZ - 186 75  
 The Czech Republic  
 krajicek@karlin.mff.cuni.cz